

DOCUMENT RESUME

ED 191 916

TM 800 540

AUTHOR Capie, William: And Others
TITLE Using Pupil Achievement to Validate Ratings of Student Teacher Performance.
PUB DATE Apr 80
NOTE 35p.: Paper presented at the Annual Meeting of the American Educational Research Association (64th, Boston, MA, April 7-11, 1980).
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Academic Achievement; *Classroom Observation Techniques; *Competency Based Teacher Education; *Correlation; Elementary Education; Higher Education; Predictor Variables; Pretests Posttests; Student Teacher Supervisors; *Student Teaching; Teacher Effectiveness; *Teacher Evaluation; Teaching Skills
IDENTIFIERS Georgia; *Teacher Performance Assessment Instruments

ABSTRACT

The magnitude of simple and multiple correlations between the Teacher Performance Assessment Instruments (TPAI) ratings of elementary student teachers and achievement gains of their pupils is assessed, as well as the impact of certain methodological and contextual factors on the validity coefficients generated. The TPAI is a set of four instruments including: Teaching Plans and Materials; Classroom Procedures; Interpersonal Skills; and Professional Standards. Participants were 33 undergraduate interns in early elementary education and several hundred pupils in grades K through 5. Interns were rated by the classroom supervising teacher and the college supervisor twice: once during the first quarter of internship and again during the second quarter. During both quarters the pupils were pre- and posttested on unit objectives developed by the interns. The pre- and posttests were used to assess student growth, and an effectiveness index (i.e., the class mean of the percentage of the possible gain) was computed for each pupil. Results indicated that 77 of 105 TPAI ratings on the first occasion, and 54 of 200 on the second occasion, were significantly and positively related to achievement gains. (RL)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED191916

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

USING PUPIL ACHIEVEMENT TO VALIDATE
RATINGS OF STUDENT TEACHER PERFORMANCE

"PERMISSION TO REPRODUCE THIS
MATERIAL ~~_____~~
HAS BEEN GRANTED BY

W. Capie

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

William Capie
Kenneth G. Tobin
Margaret Howell

University of Georgia

Athens, GA

30602

Paper presented
at the
Annual Meeting of the American Educational
Research Association

Boston, MASS
April 1980

2

TM 800540

Using Pupil Achievement to Validate Ratings of Student Teacher Performance

Teacher training has long been an area of considerable controversy and debate. The movement toward Competency or Performance based Teacher Education (C/PBTE) in the sixties was seen as a means to systematize and, hence, improve the quality of teacher training and teaching. In the wake of C/PBTE came a new set of questions and debates. Education courses were questioned on the basis of the validity of their content. Researchers responded to demands for valid competencies with reviews of previous studies and proposals for new research. Rosenshine and Furst (1971) provided direction for training and research with their review of selected studies which identified eleven of "the most promising of the variables studied." Subsequent inspection of these same studies drew a shadow over the optimism stimulated by their initial review. Heath and Nielson (1974), after careful scrutiny, questioned the validity of the variables, reliability of measurements, statistical analyses, adequacy of reporting and raised other similar issues relating to the original research. They concluded

. . . the research literature on the relation between teacher behavior and student achievement does not offer an empirical basis for the prescription of teacher training objectives.

They and others (cf Berliner 1976) imply that the task of establishing links between teacher and outcomes is so complex and confounded by other variables that any results "are likely to be inherently trivial." (Heath and Nielson, 1974)

The problems notwithstanding, state legislatures and other agencies have created a tremendous impetus for the C/PBTE concept by enacting various types of performance based certification and/or education legislation.

The state of Georgia is implementing a performance assessment program as a part of the teacher certification process. The assessment is multifaceted in that content knowledge is measured as well as in-the-classroom performance. Classroom assessments are completed during student teaching as well as during a three year probationary period. Successful performance is to be required before an individual receives professional certification. The competencies were identified through a statewide survey of some 5000 educators who rated a large number of competency statements on the scale--essential, highly desirable, desirable, or none of these. (Johnson, Okey, Capie, Ellett, and Adams, 1979) The Teacher Performance Assessment Instruments (TPAI) (Capie, Johnson, Anderson, Ellett, and Okey, 1979) have evolved over a three year period to assess competencies which were rated essential. A preliminary edition was used in this study.

For certification purposes the TPAI will be used with student teachers during internship and on several occasions with beginning teachers during a three year probationary period. The instruments have been studied extensively with both groups to determine their utility and their technical characteristics. Modifications have been made based on formal and informal feedback, studies of interrater agreement and internal consistency, factor analyses, and various validity investigations. The best measure of validity has been considered the ability to predict student achievement. Consequently, this study was planned to determine what portion of the variation in pupil achievement could be explained by student teachers' ratings on the TPAI.

Objective

The purpose of the study was to assess the magnitude of simple and multiple correlations between TPAI ratings of elementary student teachers and achievement gains of their pupils. A secondary purpose was to examine the impact of certain methodological and contextual factors on the validity coefficients generated.

Methods

Participants in the study were 33 undergraduate interns in early elementary education and several hundred pupils in public school classrooms in grades K-5.

Data Collection. The interns visited the classrooms for two days early in the first quarter of a two-quarter senior internship sequence. Subsequently they returned to campus and prepared a short science unit for use in the school. During the first full week of school activity the unit was implemented. The interns were rated on the TPAI by their science methods instructor as well as an instructor in general methods after independent observations of one instructional period. The pupils were pre-tested and post-tested on the unit objectives with instruments developed by the interns.

During the second quarter of internship (a period like student teaching) a second, longer unit in a subject area of the interns' choice was prepared. In this case the TPAI ratings were completed by the college supervisor (who had been the general methods instructor) and the classroom supervising teacher. Again pre- and post-tests were administered to the pupils.

Instruments. The TPAI is a set of four instruments which are an integrated part of the certification plan for the state of Georgia.

The Teaching Plans and Materials (TPM) instrument addresses competencies related to planning and evaluating instruction. The Classroom Procedures (CP) instrument includes competencies pertinent to instruction and classroom administration. The Interpersonal Skills (IS) instrument addresses affective elements of teaching. The Professional Standards (PS) instrument addresses professional and ethical dimensions of teaching. The TPAI edition used with student teachers addressed 17 competencies. Each was defined by 2-5 performance indicators rated on a 5 point scale representing degrees of increasing proficiency. The five scale points have been defined with 1-2 sentence descriptors in an effort to reduce ambiguity. A sample is included in Figure 1. Numerous reliability measures have been reported, (Capie, Ellett, and Johnson, 1979) including interrater agreement rates of .85-.91 and Cronbach alpha estimates on the order of .8. Tobin and Capie (1980) have reported generalizability coefficients of .97 to differentiate teachers and competencies while generalizing over raters.

CP Indicator 18. *Attends to routine tasks.*

Scale of Descriptors:

1. The teacher does not attend to routine tasks.
2. The teacher attends to routine tasks in a disruptive or inefficient manner (e.g., learners need special permission for many routine tasks).
3. The teacher anticipates routine tasks and attends to them efficiently.
4. Routine tasks are handled smoothly. Teacher delegates many tasks to the students.
5. In addition to 4, learners are responsible for various dimensions of the task (e.g., distributing materials, picking up work area, etc.).

Figure 1 A sample TPAI indicator

Many teacher performance dimensions were assessed. Thirty-six TPAI indicators comprising twelve competencies were rated on the first occasion.

On the second occasion all forty-nine indicators (comprising seventeen competencies) were rated. Competency scores are defined as the mean of all relevant indicator ratings. In addition, factor scores based on analyses of data sets from 857 beginning teachers were computed. (Tobin, Capie, Ellett and Johnson, 1980). All of these scores were available from each of the two data sources as well as composites of the two.

Pre- and post-tests constructed by the interns were used to assess student growth. The interns were asked to develop five or more performance objectives which were assessable with objective test items. A balance of cognitive levels was required. Test length was not fixed, but multiple items were encouraged for most objectives. Each test was reviewed by the investigators to ensure congruence of objectives and tests with the established ground rules and suitability for learners. Modifications in tests were made where necessary.

An effectiveness index was derived from the pre- and post-tests. The effectiveness index was the class mean of a measure called the percentage of the possible gain computed for each pupil. The difference between the post-test and the pre-test (the observed gain) was divided by the difference between the maximum possible score and the pre-test score (the possible gain) and then multiplied by 100 to compute the percentage of possible gain. The characteristics of this and other effectiveness indices have been documented elsewhere. (Safter, Tobin, and Capie, 1980).

Analysis

For each intern three performance ratings were available for each unit. Two ratings were made independently by different observers; the third was the composite (mean) of the first two. Pearson product-moment

correlations were computed between the effectiveness index for each unit and the indicator and competency ratings on each occasion.

Factor scores representing performance in four areas were also computed from each set of ratings. The four areas were General Teaching Effectiveness (heavily weighted with planning), Professional Standards, Learning Environment and Management. These factor scores were then used in a regression analysis against the effectiveness index. In each case the most parsimonious solution which accounted for a significant proportion of the variance in the criterion measure was selected. Factor scores were used in the regression analysis because of the generalizability demonstrated in the early study. (Tobin, et al., 1980) The parsimony of four variables as compared with forty-five indicators was also a consideration when the small sample size was considered in this study.

Results

On the first occasion seventy-seven of a possible one hundred five ratings were significantly and positively related to achievement gains in science. A summary table is reported in Table 1. Correlations ranged to .72 for indicators (science instructor's ratings of IS 3, "Demonstrates Warmth and Friendliness"), and to .66 for competencies (science instructor's ratings of Competency 14, "Helps Learners Develop Positive Self-Concepts"). A complete listing of indicator correlations for individual and combined data sources is included as Table 2. A complete listing of competency correlations for individual and combined data sources is included as Table 3.

(Include Tables 1, 2, and 3 About Here)

When factor scores were correlated with mean achievement gains in science significant correlations were found in each case. Values ranged from .5 to above .6 In all cases, the multiple regression improved prediction only slightly beyond that of either factor considered alone. Complete data are contained in Table 4.

(Include Table 4 About Here)

On the second rating occasion fifty-four of a possible two hundred ratings were significantly and positively related to achievement gains. A summary is included in Table 5. Correlations ranged to .50 for indicators (college supervisors ratings of TPM 5, "Plans instruction at a variety of cognitive levels"), and to .41 for competencies (college supervisors ratings of Competency 2, "Organizes instruction to take into account individual differences among learners"). A complete listing of indicator correlations for individual and combined data sources is included in Table 6. A complete listing of competency correlations for individual and combined data sources is included as Table 7.

(Include Tables 5-7 About Here)

When factor scores were correlated with achievement gain, correlations ranged to .37 (Supervising teachers' rating of factor 3. Professional Standards). A complete listing of factor score/gain correlations for the second occasion is included in Table 8.

(Include Table 8 About Here)

Only one two-predictor model was identified through multiple regression. The model including the supervising teachers' ratings of Management and General Teaching Effectiveness accounted for 22% of the variance in class mean gains.

Discussion

These two data sets along with an earlier study involving secondary student teachers (Capie, Ellett, and Johnson, 1980) are parallel studies of the TPAI and the underlying competencies. The implications of the findings for teaching and teacher training have been discussed elsewhere (Capie, Tobin, and Bowell, 1980). Each replication is somewhat different in terms of design and outcomes. Therefore, this paper will address methodological considerations in measurement and design in teacher assessment studies of this type.

Although the TPAI were used in each study the characteristics of the teaching performance measures were undoubtedly affected by many factors in the study. Seven important dimensions in these studies are:

1. Variation in Teaching Performance
2. Variation in Raters
3. Expertise of Raters
4. Quality of Tests
5. Amount of Growth in Unit
6. Curriculum Area of Study
7. Context

In studies involving the TPAI these variables may influence the magnitude of (and hence the number) of validity coefficients. Table 9 summarizes each of the three studies with regard to these seven dimensions. Each dimension is discussed in more detail.

Variation in Teaching Performance. In any correlational study variation in the measures is necessary. The elementary science study was conducted in the pre-student teaching quarter where performance levels are diverse. The secondary study occurred in four different colleges in Georgia; again representing diverse performance levels. The elementary student teaching study occurred in a program with frequent, specific, TPAI related supervision. Variation in TPAI ratings in this study was

reduced on the average of .5 standard deviation across all indicators compared to the other studies.

Variation among Raters. Raters often see a teaching performance in different ways. Consequently they contribute a certain bias to a data set. The effect could be statistically eliminated in a situation where every rater observed every teacher, but it is difficult to deal with in other contexts. In the elementary science study the same two raters did observe and rate all teachers. In the elementary student teacher study one of the original raters again observed and rated all the student teachers. However, in this study a different supervising teacher rated each of the student teachers. The secondary study was conducted with the two raters nested within teachers; that is each teacher had a different college supervisor as well as a different supervising teacher.

Expertise of Raters. Errors can exist within raters as well as between them. Training would tend to minimize both sources of variation. The science study was conducted by two individuals who had been heavily involved in statewide training of observers. Although one of these well-trained raters participated in the elementary student teacher study, the classroom supervisors were not trained in the use of the TPAI. In contrast, the supervising teachers in the secondary study were trained according to state criteria but the college supervisors were not.

Quality of Tests. The tests in each of these studies were constructed by the student teachers. Consequently, they are suspect on grounds of quality. In both elementary studies the tests were reviewed by the investigators to ensure their suitability for learners, objectives and the research purpose. Tests were modified when necessary. Students in the secondary study were merely asked to have their tests match the stated criteria to the greatest extent possible; but, no monitoring or revision was possible.

Amount of growth in Unit. Just as variation in teacher performance is necessary to obtain large correlations, so is variation in the dependent measure. The science study represented more variation in the dependent measure than did either of the others.

Curriculum Area of the Study. The curriculum area is one of a number of context variables that may be important. The science study, of course, represented a number of similar unit topics. All classes had a similar "process" approach with a focus on the manipulation of quantitative data. In both other studies, the content of the unit was not controlled.

Context of Study. Both elementary studies occurred in two schools located within similar neighborhoods of the same community. The secondary study sample was drawn from four colleges with student teachers placed in many different communities throughout the state. While the gain score used as an effectiveness index provides for different entry abilities, community or class differences could still be an important factor in confounding results.

Interpretation

Reviewing the seven factors suggests that the strong controls and diverse performance levels occurring in the elementary science study would suggest the greatest potential for identifying a valid relationship. That prediction was true in that the number and magnitude of significant correlations was greater in that study than in either of the others. Apparently the TPAI do discriminate among less able student teachers when they exist in the sample. A data summary is included in Table 9.

The other two studies are difficult to characterize because the degree of control varies as does the performance levels of the student teachers.

Interestingly, the ratings of the college supervisor which produced significant validity coefficients during the science study failed to do so during the second study. The level of performance and the reduced range of gain scores may have suppressed the correlations.

While the secondary study was characterized by poor controls, significant validity coefficients were obtained when the ratings of the classroom supervisors were used as predictors. The training of these individuals may have accounted for the greater number of significant correlations when compared to the untrained college observers. The secondary study was similar to other studies which have been conducted using the TPAI with beginning teachers. (Ellett, Capie, and Johnson, 1980 a,b) These studies have tended to show low to moderate correlations for all indicators and competencies when trained data collectors are used to make the observations regardless of the mix of grade and subject areas represented in the sample.

Conclusions

The difficulty in documenting the validity of competencies, instruments or specific teaching performances is well known. TPAI are unusual in that several replications of studies have shown that they consistently predict student achievement gains. This paper has illustrated that even when valid instruments are used, methodological or contextual factors can enhance the results or mask validity which may exist. A key factor in correlational studies such as these is identifying a group of teachers which is characterized by a range of performance. This task is not easy given the selection and training that occur as candidates progress through their preservice programs. A second, major concern is assuming that accurate and reliable performance measures be obtained.

Two types of studies seem necessary as process product studies are proposed and implemented. On the one hand, tightly controlled studies are desirable where rater effects, curriculum differences, etc. are minimized. On the other hand, studies without the controls are also necessary to illustrate the representativeness of the findings. These representative studies should be done with somewhat larger samples to overcome the power loss due to reduced reliability of the measures obtained.

TABLE 1

Correlations Between Teacher Ratings
and Pupil Achievement Gains in Science

	<u>Science Instructor</u>	<u>College Supervisor</u>	<u>Combined Ratings</u>
Number of Indicators Rated	31	32	33
Number of Significant Correlations	17	13	19
Number of Competencies Rated	11	11	11
Number of Significant Correlations	9	4	9
Number of Factors Rated	2	2	2
Number of Significant Correlations	2	2	2

Table 2

Correlations of Elementary Student Teachers TPAI Ratings with
Student Achievement Gains in a Science Unit (25 classrooms)

<u>Indicators</u>		<u>Prime¹ Factor</u>	<u>College Supervisor</u>	<u>Science Instructor</u>	<u>Combined Ratings</u>
TPM 1	Specifies or selects learner objectives for lessons.	G	.07	.12	.12
TPM 2	Specifies or selects teaching procedures for lessons.	G	.13	-.03	.05
TPM 3	Specifies or selects content, materials, and media for lessons.	G	-.08	.25	.15
TPM 4	Specifies or selects materials and plans for assessing learners.	G	-.01	.14	.05
TPM 5	Plans instruction at a variety of cognitive levels.	G	.12	.21	.21
CP 1	Uses audiovisual and other instructional equipment.	G	-.23	.64****+	.42+++
CP 2	Uses instructional materials that provide appropriate practice on the objectives.	G	.30	.38*	.42*
CP 3	Organizes learning activities in a logical sequence.	M	.37*	.14	.29
CP 4	Gives clear directions and explanations.	M	.23	.23	.29
CP 5	Provides directions and explanations when learners misunderstand.	M	.37*	.33*	.42*

Table 2 (Continued)

	<u>Indicators</u>	<u>Prime¹ Factor</u>	<u>College Supervisor</u>	<u>Science Instructor</u>	<u>Combined Ratings</u>
CP6	Conveys the impression of knowing what to do and how to do it.	L	.16	.35*	.29
CP 7	Uses acceptable written and oral expression with learners.	L	.18	.33	.31
CP 8	Uses responses and questions from learners in teaching.	L	.28	.44*	.45**
CP 9	Provides learners with opportunities for participating.	M	-.07	.17	.06
CP 10	Demonstrates ability to conduct lessons using a variety of teaching methods.	L	.21	.66***	.56**
CP 11	Uses teaching methods appropriate for objectives and learners.	M	.05	.38*	.25
CP 12	Demonstrates ability to work with individuals, small groups, and large groups	M	.23	.18	.27
CP 13	Uses procedures which initially get learners involved in lessons.	L	.38*	.42*	.48**
CP 14	Maintains learner involvement in lessons.	M	.51**	.58**	.62***
CP 15	Demonstrates knowledge in the subject area.	L	.41*	.25	.38*
CP 16	Presents accurate information about the topic being taught.	L	.39*	.16	.26

Table 2 (Continued)

	<u>Indicators</u>	<u>Prime¹ Factor</u>	<u>College Supervisor</u>	<u>Science Instructor</u>	<u>Combined Ratings</u>
CP 17	Provides learners with information on the purpose and importance of topics.	L	.33*	.23	.33*
CP 18	Attends to routine tasks.	M	.49**	.52**	.55**
CP 19	Organizes materials for efficient use by learners.	M	.39*	.45*	.48**
CP 20	Provides a learning environment that is attractive and orderly.	L	.79*+	--	.76*+
CP 21	Uses special or unexpected events to supplement lessons.	M	--	--	--
CP 22	Makes modifications in lessons as needed.	M	--	--	.89**++++
IS 1	Communicates personal enthusiasm.	L	.10	.26	.19
IS 2	Communicates with learners in a way that conveys interest in them and in the subject.	L	.38*	.50**	.53*
IS 3	Demonstrates warmth and friendliness.	L	.23	.72***	.60***
IS 4	Demonstrates sensitivity to the needs and feelings of learners.	L	.24	.40*	.49**
IS 5	Demonstrates patience, empathy, and understanding.	L	--	--	--
IS 6	Provides feedback to learners about their behaviors.	M	.60***	.05	.56**

Table 2 (Continued)

	<u>Indicators</u>	<u>Prime¹ Factor</u>	<u>College Supervisor</u>	<u>Science Instructor</u>	<u>Combined Ratings</u>
IS 7	Promotes comfortable inter- personal relationships.	M	.16	.40*	.30*
IS 8	Manages disruptive behavior among learners.	M	.63***	.45**	.59***
IS 9	Maintains appropriate classroom behavior.	M	.33	.53**	.53**

¹ Factor Designations are as follows: G = General Teaching Effectiveness; L = Learning Environment;
M = Classroom Management.

*p < .05
**p < .01
***p < .001

+N = 10
++N = 12
+++N = 16
++++N = 6

TABLE 3

Correlations of Elementary Teachers TPAI Competency Ratings with
Student Achievement Gains in a Science Unit (25 classrooms)

<u>Competency</u>	<u>College Supervisor</u>	<u>Science Instructor</u>	<u>Combined Ratings</u>
1. Plans instruction to achieve selected objectives.	.06	.17	.14
2. Organizes instruction to take into account individual differences among learners.	--	--	--
3. Uses instructional techniques, methods, and media related to the objectives.	.32	.36*	.41*
4. Communicates with learners both orally and in writing.	.29	.42*	.41*
5. Reinforces and encourages the efforts of learners.	.14	.34*	.30
6. Demonstrates a repertoire of teaching methods.	.24	.52**	.47**
7. Uses procedures which involve the learner in the instruction.	.51**	.56**	.62***
8. Demonstrates an understanding of the school subject being taught.	.55**	.21	.42*
9. Organizes time, space, materials and equipment for instruction.	.51**	.56**	.61***
10. Adjusts instruction to changes in conditions.	-.50 ⁺	--	.86***+

TABLE 3 (Continued)

<u>Competency</u>	<u>College Supervisor</u>	<u>Science Instructor</u>	<u>Combined Ratings</u>
11. Demonstrates enthusiasm for teaching and learning and the subject being taught.	.30	.50**	.45**
12. Helps learners develop positive concepts of themselves.	.27	.66***	.61***
13. Manages classroom interactions.	.50**	.49**	.59**
14. Obtains and uses information about the needs and progress of individual learners.	--	--	--
15. Obtains and uses information about the effectiveness of instruction to revise it when necessary.	--	--	--

*p $\leq$.05

**p $\leq$.01

***p $\leq$.001

+N = 6

++N = 7

TABLE 4

Correlation of Factor Scores
and Mean Achievement Gains

<u>Factor</u>	<u>Science Instructor</u>	<u>College Supervisor</u>	<u>Composite</u>
Management	.61	.50	.61
Learning Environment	.62	.48	.62
Multiple R	.63	.50	.63

TABLE 5

Correlations Between Student Teacher Ratings
and Pupil Achievement Gains

	<u>College Supervisor</u>	<u>Supervising Teacher</u>	<u>Combined Rating</u>
Number of Indicators Rated	46	47	47
Number of Significant Correlations	13	9	16
Number of Competencies Rated	16	16	16
Number of Significant Correlations	4	5	3
Number of Factors Rated	4	4	4
Number of Significant Correlations	1	2	1

Table 6

Correlations of TPAI Ratings of Student Teachers
With Mean Pupil Achievement Gains (30 Classrooms)

TPM Indicator	College Supervisor	Classroom Supervisor	Combined Ratings
1	.26	-.14	.08
2	.32*	.22	.32*
3	.04	.12	.10
4	.04	-.09	.03
5	.10	.50*	.46**
6	.32*	.05	.17
7	.34*	.16	.24
8	.31*	.14	.27
9	.31*	-.17	.00
10	.05	.05	-.01
11	-.02	.14	.00
12	.24	.25	.30*

*p < .05

**p < .01

Table 6 Continued

Correlations of TPAI Ratings of Student Teachers
With Mean Pupil Achievement Gains (30 Classrooms)

CP Indicator	Colleges Supervisor	Classroom Supervisor	Combined Ratings
1	-.03	-.01	.01
2	-.14	.06	-.05
3	-.00	.12	.09
4	.28	.24	.29
5	.22	.02	.13
6	.22	.00	.12
7	-.03	.24	.09
8	.15	.22	.22
9	.09	.35*	.25
10	.11	.28	.24
11	.40*	.15	.36*
12	.00	.20	.12
13	.14	.18	.18
14	.27	.25	.31*
15	.29	-.06	.13
16	.27	.01	.14
17	.47**	.11	.35*
18	.12	.41*	.36*
19	.01	.22	.12
20	.49**	.20	.34*
21			
22			

*p $\leq$.05**p $\leq$.01

Table 6 Continued

Correlations of TPAI Ratings of Student Teachers
With Mean Pupil Achievement Gains (30 Classrooms)

IS Indicators	College Supervisor	Classroom Supervisor	Combined Ratings
1	.19	.27	.26
2	.32*	.20	.30*
3	.11	.22	.30
4	.32*	.32*	.36*
5		.10	.14
6	.20	.39*	.33*
7	.36*	.39*	.40*
8	.39*	.16	.30*
9	.33*	.31*	.37*
PS Indicators			
1	.07	.32*	.19
2	-.33*	.25	.01
3	.11	.28	.20
4	.21	.22	.24
5	.16	.36*	.31*
6	.21	.26	.27

*p $\leq$.05**p $\leq$.01

Table 7

Correlations of TPAI Ratings of Student Teachers
With Mean Pupil Achievement Gains (30 Classrooms)

Competency	College Supervisor	Classroom Supervisor	Combined Ratings
1	.19	.17	.22
2	.41*	.13	.21
3	.21	-.06	.01
4	.16	.20	.20
5	-.11	.07	.00
6	.23	.09	.19
7	.14	.33*	.27
8	.25	.25	.28
9	.23	.27	.27
10	.40*	.01	.22
11	.18	.34*	.31*
12	.62+		
13	.30*	.25	.29
14	.24	.23	.27
15	.38*	.34*	.39*
16	.09	.37*	.24
17	.19	.32*	.30*

*p < .05
+n = 6

TABLE 8

Correlations Between Teacher Factor Scores and Pupil
Achievement Gains on the Second Occasion

Factor	<u>College Supervisor</u>	<u>Supervising Teacher</u>	<u>Combined Ratings</u>
Management	.29	.32	.33
General Teaching Effectiveness*	.27	.13	.18
Professional Standards	.14	.37	.28
Learning Environment	.33	.22	.29

*Planning Skills on GTE.

Table 9

Context Variables Related To Process
Product Validity Coefficients

Study	Elementary Science		Elementary Student Teacher		Secondary Student Teacher	
	Science Instructor	College Supervisor	Supervisor Teacher	College Supervisor	Supervisory Teacher	College Supervisor
Teaching Performance	Great Variation		Little Variation		Great Variation	
Rater Variation	None	None	1 rater per teacher	None	1 rater per teacher	1 rater per teacher
Rater Expertise	High	High	Low	High	Satisfactory	Low
Test Quality	High		High		Unknown	
Amount of Pupil Growth	Great Variation		Less Variation		Great Variation	
Curriculum Area	Controlled		Variation		Great Variation	
Context	Controlled		Controlled		Great Variation	
Percent Signi- ficant Corre- lations	55	41	19	28	41	14
Magnitude of Factor Corre- lations	.6	.5	.3	.3	.2	.11

References

- Berliner, D. "Impediments to the Study of Teacher Effectiveness," Journal of Teacher Education, 1976, 47:5, 13.
- Capie, W., Johnson, C., Anderson, S.J., Ellett, C., and Okey, J. Teacher Performance Assessment Instruments. Athens, Georgia, Teacher Assessment Project, 1979.
- Capie, W., Ellett, C., and Johnson, C. Selected Investigations of the Reliability of the TPAI, Athens, Georgia: Teacher Assessment Project, 1979.
- Capie, W., Ellett, C., and Johnson, C. Relating Pupil Achievement Gains to Ratings of Secondary Student Teachers Performance, Paper presented at the Eastern Educational Research Association, Norfolk, VA., 1980.
- Capie, W., Tobin, K., and Bowell, M. Using Science Achievement to Validate Student Teacher Competencies, Paper presented at the Annual Meeting of the National Association for Research in Science Teaching, Boston, MA., April 1980.
- Ellett, C., Capie, W., and Johnson, C. Teacher Performance and Pupil Perceptions of the Social Environment of Learning. Paper presented at Eastern Educational Research Association, Norfolk, VA., March 1980A.
- Ellett, C., Capie, W., and Johnson, C. Teacher Performance and Pupil Achievement on Teacher Made Tests, Paper presented at Eastern Educational Research Association, Norfolk, VA., 1980b.
- Heath, R. W., and Neilson, M. A. "The Research Base of Performance Based Teacher Education," Review of Educational Research, 1974, 44:463-84.
- Johnson, C., Okey, J., Capie, W., Ellett, C., and Adams, P. Identifying and Verifying Generic Teacher Competencies. Athens, Georgia: Teacher Assessment Project, 1978.
- Rosenshine, B. V., and Furst, N. "Research on Teacher Performance Criteria." In B. O. Smith (Ed.) Research in Teacher Education: A Symposium. Englewood Cliff, N.J. Prentice Hall, 1971.
- Safer, W., Tobin, K. and Capie, W. An Empirical Analysis of Teacher Effectiveness Indices Based on Class Mean Change Scores. Paper presented at Eastern Educational Research Association, Norfolk, VA., 1980.
- Tobin, K. and Capie, W. Applications of Generalizability Theory to Teacher Performance. Paper presented at Southeastern Invitational Conference on Measurement in Education, Greensboro, February, 1980.
- Tobin, K., Capie, W., Ellett, C., and Johnson C. A Confirmatory Analysis of the Factor Structure of Teacher Performance. Paper presented at Eastern Educational Research Association, Norfolk, VA., 1980.